BP neural network classification on passenger vehicle type based on GA of feature selection

QIN Hui-chao (秦慧超), HU Hong-ping (胡红萍), BAI Yan-ping (白艳萍)
(*School of Science, North University of China, Taiyuan 030051, China*)

Abstract: This paper has concluded six features that belong to passenger vehicle types based on genetic algorithm(GA) of feature selection. We have obtained an optimal feature subset, including length, ratio of width and length, and ratio of height and length. And then we apply this optimal feature subset as well as another feature set, containing length, width and height, to the network input. Back-propagation(BP) neural network and support vector machine(SVM) are applied to classify the passenger vehicle type. There are four passenger vehicle types. This paper selects 400 samples of passenger vehicles, among which 320 samples are used as training set (each class has 80 samples) and the other 80 samples as testing set, taking the feature of the samples as network input and taking four passenger vehicle types as output. For the test, we have applied BP neural network to choose the optimal feature subset as network input, and the results show that the total classification accuracy rate can reach 96%, and the classification accuracy rate of first type can reach 100%. In this condition, we obtain a conclusion that this algorithm is better than the traditional ones[9].

Key words: genetic algorithm(GA); feature selection; back-propagation(BP) network; passenger vehicles type

CLD number: TP183 **Document code**: A
Article ID: 1674-8042(2012)03-0251-04 **doi**: 10.3969/j.issn.1674-8042.2012.03.011

0 Introduction

Electronic toll collection (ETC) is an important part of intelligent transport system (ITS). In recent years, It has been widely used in the world, and China also pays great attention to it. Many provinces and cities in China are transforming the toll way with the application of ETC. In addition, some areas have put it to practice. ETC has advantages in certain aspects, such as energy saving, emission reduction, the increasing of passing rate and the reducing of working strength and so on. Toll collection is based on the type of vehicle. In this condition, although vehicle information has been stored in the integrated circuit(IC) card, there is still the existence of leaking tolls by swapping IC card. Combining automatic vehicle type classification with IC card can effectively avoid tolls' loss. This paper applies genetic algorithm(GA) feature selection to choose six features about passenger vehicle types for an optimal subset, and uses back-propagation(BP) neural network to classify four types of passenger vehicles for better results.

Artificial neural network (ANN) is actually a complex network which is connected by simple components. The specific ANN consists of many neuron models which are the same, parallel and interconnected. And the signal is processed by the interactions in neurons[1].

Genetic algorithm is an adaptive probability optimization technology based on genetic and evolutionary mechanism for the optimization of complex systems, and it is a random search algorithm which learns from natural selection natural genetic mechanisms[2].

1 Models about BP and SVM

BP network is the kernel of the feed-forward neural network. 80% − 90% of the ANN models use BP network or its changing forms. BP network is a multilayer network which conducts weights training of nonlinear differentiable function. BP algorithm consists of information forward propagation and error back propagation. In the error back propagation, the error is returned along the original pathway in the network to modify the weights of the neuron on each layer until it reaches the expected value. BP network structure with a hidden layer is shown in Fig. 1.

* **Received data**: 2012-03-22
Foundation item: China Postdoctoral Science Foundation (No. 20100481307), Natural Science Foundation of Shanxi(No. 2009011018-3)
Corresponding author: QIN Hui-chao (qinhuichao2@163.com)

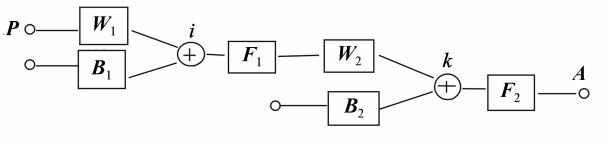

Fig. 1 BP neural net model with a hidden layer

In Fig. 1, $\boldsymbol{P}$ denotes the input, $\boldsymbol{A}$ denotes the predicted output, i is the number of hidden layer neurons, k the number of output neurons, $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ denote the weights, $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ denote the biases, $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ denote the activation transfer functions. The number of input neurons equals to the number of matrix $\boldsymbol{P}$'s rows. The formula of the network output can be described as

$$\boldsymbol{A} = \boldsymbol{F}(\boldsymbol{WP} + \boldsymbol{B}), \tag{1}$$

where $\boldsymbol{A}, \boldsymbol{F}, \boldsymbol{W}, \boldsymbol{P}$ and $\boldsymbol{B}$ stand for matrix.

Activation transfer function commonly uses threshold function, sigmoid and linear function.

Support vector machine (SVM) is an algorithm which can make the input vector be mapped into a high-dimensional featured space through a pre-selected nonlinear mapping. An optimal classification super plane can be built in this space in order to transform low-dimensional ones being linearly inseparable to high-dimensional samples being linearly separable. Such transformation can be realized with the application of the kernel functions[3]. The original SVM is a binary classifier. SVM schematic diagram is shown in Fig. 2.

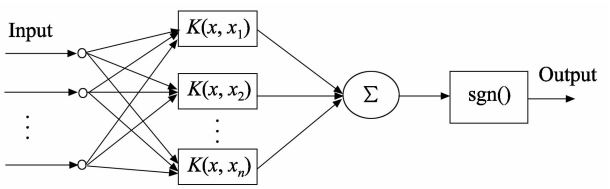

Fig. 2 SVM schematic diagram

Radial basis function (RBF) kernel function is a commonly used one which is shown as

$$K(\boldsymbol{x}, \boldsymbol{x}_i) = \exp(- \| \boldsymbol{x} - \boldsymbol{x}_i \|^2 / 2\delta^2), \tag{2}$$

where, $\boldsymbol{x}$ and $\boldsymbol{x}_i$ denote feature vectors of the sample.

2 Using GA for feature selection

Feature selection is the process to select a group of optimal feature subset $\boldsymbol{d}$ ($\boldsymbol{d} < \boldsymbol{D}$)) from feature set $\boldsymbol{D}$[4]. Feature selection can help diminish the irrelevant or redundant features so as to achieve reduction of feature number, improvement of accuracy of the model and reduction of running time.

The process of feature selection generally includes generation procedure, evaluation function, stopping criterion and validation procedure. Stopping criterion is related to the evaluation function, which generally refers to a threshold. The search stops when the value of the evaluation function is in the threshold. Validation procedure can verify the validity of the feature subset based on the data set. The process of feature selection is shown in Fig. 3.

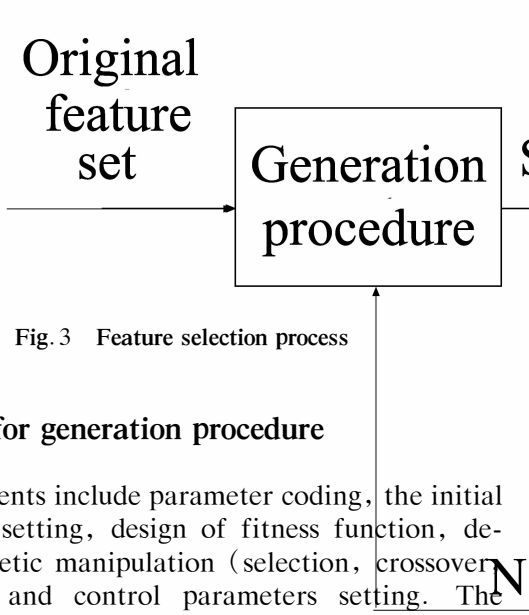

Fig. 3 Feature selection process

2.1 GA for generation procedure

GA elements include parameter coding, the initial population setting, design of fitness function, design of genetic manipulation (selection, crossover, mutation) and control parameters setting. The higher the fitness value is, the better the individual is. Crossover is a main method to generate a new individual. Crossover operator and mutation operator cooperate together to complete the global search and local search in the search space.

Generation procedure is a process to search feature subspace. Searching algorithms can be divided into complete algorithm, heuristic algorithm and random algorithm. GA belongs to random algorithm[5,6].

The parameter coding is applied in this paper, in which the chromosome length is 6, and each gene represents a feature. Allele value is {0,1}, where 0 means that we should not select this feature and 1 means the opposite[7].

A typical general population of GA is composed of 30 – 100 individuals. However, due to the limited number of the feature set of this paper, we select the population with 10 individuals. The fitness function is

$$G(\boldsymbol{A}) = \frac{1}{(\boldsymbol{T} - \boldsymbol{A})^2}, \tag{3}$$

where $\boldsymbol{T}$ denotes the desired output and $\boldsymbol{A}$ is the note of the predicted output.

Selection is an operation of choosing the fittest individuals. The selection function in this paper utilizes the method of stochastic universal sampling, which chooses an individual according to the breeding probability transformed from the fitness of the

individual in the current population, and then the selected individual is transferred to the next population. This paper uses a single point crossover operator and a discrete mutation operator for the operation of crossover and mutation. Control parameters are the optional parameters.

2.2 Evaluation function

Evaluation function is employed to evaluate the feature subset provided during generation procedure. In terms of the working principle, the evaluation function can be divided into filter and wrapper[8]. In this paper, the wrapper is used as the evaluation function, and essentially it is a classifier. To classify the sample set, wrapper uses the selected feature subset, and the classification accuracy is to standardize the fitness of the feature subset. BP network is selected as the evaluation function in this paper, and the smaller $(\boldsymbol{T}-\boldsymbol{A})^2$ is, the higher the classification accuracy is.

3 Model and results analyze

This paper classifies the passenger vehicles into four types based on Chinese highway passenger fee scale. Classification standard is shown in Table 1.

Table 1 Classification of the passenger fee scale

Classification	⩽7 seats	8－19 seats	20－39 seats	⩾40 seats
Type	First type	Second type	Third type	Forth type

In this paper, passenger vehicle uses length, width, height and its ratios as features; six features are shown in Table 2.

Table 2 Six features of the selected vehicles

1	2	3	4	5	6
length	width	height	width/length	height/length	width/height

This paper uses GA for feature selection to select the six features, BP network as the evaluation function. In this paper, BP neural network is of multiple input and single output. The optimal result of feature selection is shown in Table 3.

Table 3 Optimal feature subset

1	4	5
length	width/length	height/length

Table 4 Four comparison results

	Optimal feature of BP	LWH of BP	Optimal feature of SVM	LWH of SVM
First type accuracy/%	100	100	100	100
Second type accuracy/%	80	70	70	80
Third type accuracy/%	95.24	95.24	85.71	85.71
Forth type accuracy/%	100	96.43	96.43	100
Total accuracy/%	96.25	93.75	91.25	93.75
Classification figure	Fig.4	Fig.5	Fig.6	Fig.7

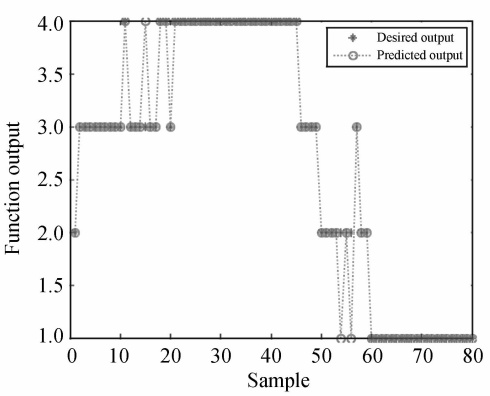

Fig.4 Optimal feature of BP

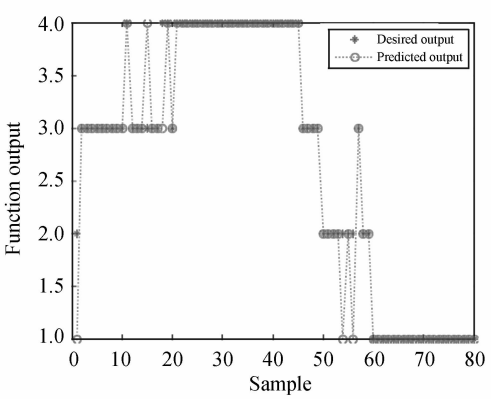

Fig.5 LWH of BP

Taking the sample features as input vectors, the vehicle types as output vectors, the type of each sample is known. When classifying, the type of vehicle as the output, where 1,2,3 and 4 represent the output. This paper uses Matlab to conduct programs and the data are normalized, then we conduct classification and prediction. This paper uses 400 samples, where 320 samples are used for training (each class contains 80 samples), and the other 80 samples for testing (the first type contains 21 samples, the second type contains 10 samples, the third type contains 21 samples, the forth type contains 28

samples).

The feature of the optimal feature subset in BP neural network is used as the feature vector. Comparing the output of BP neural network and the output of SVM[9] with three feature vectors of length, width and height(LWH), we can obtain the results listed in Table 4 and illustrated in Figs. 4 – 7.

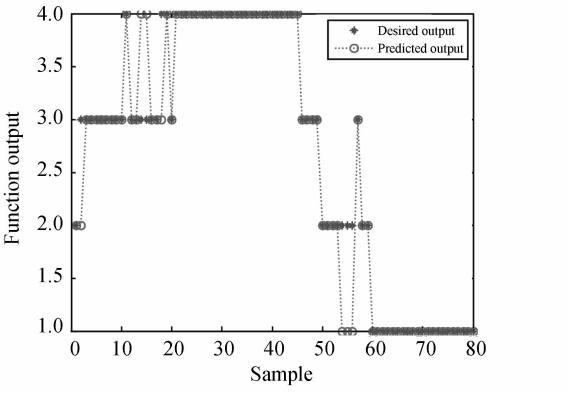

Fig. 6 Optimal feature of SVM

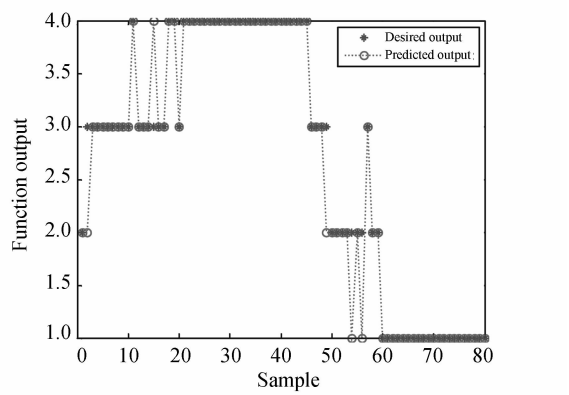

Fig. 7 LWH of SVM

4 Conclusion

ETC has become the trend of China's highway toll for its development, and a good classification can contribute to improving the working efficiency. The test results show that with the application of GA for feature selection, we can obtain an optimal feature subset which leads to an ideal classification for vehicle types.

References

[1] CONG Shuang. Neural network theory and applications with matlab toolboxes. University of Science and Technology of China Press, 2009.

[2] LEI Ying-jie, ZHANG Shan-wen, LI Ji-wu, et al. Matlab genetic algorithm toolbox and its application. Xi'an Electronic and Science University Press, 2005.

[3] Cristianini N, Taylor J S. An introduction to support vector machines and other kernel-based learning methods. Cambridge University Press, New York, 2003.

[4] MAO Yong, ZHOU Xiao-bo, ZHENG Xia. A survey for study of feature selection algorithms. Pattern Recognition and Artificial Intelligence, 2007, 20(2):211-218.

[5] Hajnayeb A, Ghasemloonia A, Khadem S E, et al. Application and comparison of an ANN-based feature selection method and the genetic algorithm in gearbox fault diagnosis. Expert Systems with Application, 2011, 38(8): 10205-10209.

[6] ZHAO Ming-yuan, FU Chong, JI Lu-ping, et al. Feature selection and parameter optimization for support vector machines: A new approach based on genetic algorithm with feature chromosomes. Expert Systems with Application, 2011, 38(5): 5197-5204.

[7] SHI Feng, WANG Xiao-chuan, YU Lei, et al. Matlab neural network of 30 cases analysis. Beihang University Press, 2010: 243-247.

[8] Foithong S, Pinngern O, Attachoo B, et al. Feature subset selection wrapper based on mutual information and rough sets. Expert Systems with Application, 2012, 39 (1): 574-584.

[9] CAO Jie, LI Hao-ru, CHEN Ji-kai. Design of vehicle type classification based on support vector machine. Science Technology and Engineering, 2007, 7(22): 5962-5965.